

Single-Case Experimental Designs

Uses in Applied Clinical Research

David H. Barlow, PhD, Michel Hersen, PhD, Jackson, Miss

The difficulty of developing and evaluating effective treatments in psychiatry and clinical psychology points out the inadequacy of current research methodology involving comparisons of large groups. An alternative approach firmly founded in the scientific method, but particularly appropriate to the study of complex behavior disorders, is the single-case experimental design. In this paper, examples of different single-case designs actually employed in applied clinical research are presented and discussed. Practical problems arising during the course of research are highlighted and some basic procedures outlined. General questions on variability, representativeness of findings, and clinical versus statistical significance are briefly discussed.

A major stumbling block to the development and evaluation of therapeutic techniques in psychiatry and clinical psychology has been the difficulty in executing meaningful clinical research. The traditional experimental group, control group, statistical analysis method of doing research has not lent itself readily to psychiatry. Matching large groups of patients with similar symptomatology is often impossible, even if one can afford the staggering costs of gathering data, following subjects, paying experimental therapists, and analyzing the data.¹ The ethical considerations of withholding treatment from control group patients, even if they eventually receive treatment, have also, rightly or wrongly, inhibited serious research.

An alternative approach, particularly appropriate for psychiatric research, is the single-case experimental design. Emanating from the case study method of psychoanalysis, on the one hand, and the laboratories of experimental (operant) psychology on the other, this approach was probably first applied to clinical problems by Shapiro.² Since then, theoretical and logical aspects of this research strategy have been discussed by Sidman,³ Dukes,⁴ Chas-

san,⁵ Baer et al.,⁶ and Davidson and Costello.⁷ Research articles employing the strategy are increasing in various psychiatric and psychological journals.

In addition to the economic and ethical issues noted above, the single-case design has several major advantages for applied clinical research whether the variables under study are psychotropic drugs or interpersonal processes.

Generality of Findings.—The first advantage is concerned with generality of findings. If an experimental group of 50 patients does statistically better than a control group of 50 patients, such differences could be due to a small number of patients in the experimental group showing larger changes while the majority of the patients show no changes or perhaps deteriorate slightly.⁸ These individual variations are masked in the group average. Furthermore, as Chasson⁹ points out, the group-statistical design does not permit conclusions as to particular patient characteristics correlated with improvement or deterioration. These data also are lost in the statistical analysis. Such findings, therefore, are not readily translatable to the practicing clinician. In the single-case design, where each patient serves as his own control in a separate experiment, effective treatments can be linked with specific patient characteristics that are immediately relevant to the clinician.

Clinical vs Statistical Change.—In a group design a treatment "works" if it produces a statistically greater effect than a control procedure. While this type of finding is very important in basic research for theoretical reasons, in clinical research these changes in patients may be so small as to be statistically significant but clinically useless. In single-case designs the size of the behavioral change in given patients is easily observed, facilitating judgments on clinical utility. Although data from single-case designs can be analyzed statistically if results are weak or unclear,¹⁰ in practice this is seldom necessary.

Mechanisms of Therapeutic Change.—Group experimental designs often test vaguely defined global treatments against no treatment, particularly in psychotherapy studies. This strategy obviates an analysis of specific mecha-

Accepted for publication Jan 8, 1973.

From the University of Mississippi Medical Center (Drs. Barlow and Hersen), and the Veterans Administration Center (Dr. Hersen), Jackson, Miss.

Reprint requests to Department of Psychiatry, University of Mississippi Medical Center, Jackson 39216 (Dr. Barlow).

nisms of change or "active ingredients" in a therapy which can then be combined with other active ingredients and used effectively by clinicians on specific patients. Single-case designs are particularly well suited for teasing out therapeutically active ingredients in a composite treatment variable.

Variability.—Finally, in group designs the effectiveness of a given treatment is usually assessed just once after treatment is completed. This strategy precludes an experimental analysis of the patient's course during treatment which, as every clinician knows, may vary considerably from day to day. A single-case design, where measures are continually taken, allows the clinical researcher to observe variability and to hypothesize which correlated environmental or personality variables may be active. New hypotheses may then be subjected to an immediate experimental analysis.

The purpose of the present paper, then, is to review single-case designs that have been employed in clinical research while providing examples of their use. Methodological and practical problems that arise during the execution of such research will be considered. Moreover, typical solutions to these problems will be outlined.

General Considerations in Single-Case Designs

A large variety of simple and complex experimental single-case designs have been developed by clinical researchers examining the effects of an equally large variety of therapeutic variables in decreasing psychopathology. Despite the differences among designs there are some common features in each. As in group designs, target behaviors (whether motor, physiological, or attitudinal) are clearly specified and methods of measurement are precisely defined. Unlike group designs, however, continuous measurements are taken throughout each phase of the study. In most designs there is an initial period of observation (baseline phase) in which the natural frequency of occurrence of the specific behavior is obtained (eg, emission of tics per unit time in a ticqueur, or scores on a depressive scale for several days prior to therapeutic intervention). Baseline measurements are generally continued until a stable pattern emerges. A minimum of three separate observation points, plotted on the graph, during this baseline phase are required to establish a trend in the data. The most desirable trend is a steady rate of behavior (eg, number of tics per minute) so that the effects of treatment, either beneficial, detrimental, or no effect, will be clear. A second trend in baseline which is common and acceptable in clinical research is one in which the patient is getting worse, a process that may have been going on for some time. With a deteriorating baseline, beneficial effects of treatment (as well as no effects) are clear. Detrimental effects, however, are less likely to be identified when behavior in baseline is already deteriorating. For the sake of convenience, the baseline phase of all designs will be labeled as the A phase throughout this paper.

In most designs, a therapeutic variable or treatment is introduced following establishment of baseline trends. Introduction of the therapeutic variable, then, represents the B phase of the experiment. Here too, a minimum of three separate observation points, and often more, are re-

quired to determine if the treatment is effective or not and whether the effect is beneficial or detrimental. In single-case, as in group, design it is most desirable to employ blind evaluation of results. In single-case designs, this means that therapists are not aware of the data during the experimental phases and those collecting the data do not know which phase of treatment the patient is undergoing.

A-B Design

The A-B experimental single-case design, briefly alluded to above, is the most basic of all designs. It represents a definite improvement over the uncontrolled case study or case history in that the target behavior is measured, and the effect of the introduction of a therapeutic variable can be determined by comparison with measured baseline rates of the behavior. However, the A-B design must be classified as a correlational design inasmuch as mere institution of the B phase does not permit unequivocal conclusions as to the controlling effects of that therapeutic variable. More specifically, changes brought about as a consequence of introducing the B variable may possibly result from its correlates rather than from its controlling effects.

An excellent example of the simple A-B design is presented by Leitenberg et al.¹⁰ They examined the effects of selective positive reinforcement on caloric intake and weight in an anorexia nervosa patient. During baseline (A phase) the patient was given four meals daily, each consisting of 1,000 calories. The patient was allowed 30 minutes per meal and was instructed that eating in a structured situation (a special hospital room) would lead to improvement. The aforementioned conditions were maintained during the B phase, but reinforcement was added. Reinforcement consisted of social praise contingent upon increasing consumption of food. In addition, privileges were made contingent upon increased consumption. Examination of Fig 1 reveals that weight maintained relative stability while caloric intake decreased slightly throughout the 50-day baseline phase (A). Institution of reinforcement (B) resulted in a marked linear increase of both caloric intake and weight over the 50-day period. Although the data suggest effectiveness of the reinforcement (eg, attention-placebo, expectancy, time, etc), only by removing reinforcement (a return to A), while holding other correlational therapeutic variables constant and noting decrement in target responses (calories and weight), is it possible to claim unequivocally that the reinforcement technique was the *sole* responsible agent of change. This latter design is known as the A-B-A experimental single-case design.

A-B-A Design

Hersen et al.¹¹ used an A-B-A design in assessing effects of a general work token economy on neurotic depression. Points earned and behavioral ratings of depression (high ratings indicate low depression) were the two target behaviors under study. Examination of Fig 2 for one neurotic depressive reveals relatively stable measurements in baseline, with a slightly upward trend for points earned and a slightly downward trend for behavioral ratings. Institution of token economy in phase B led to a dramatic upward trend in points earned and behavioral ratings. Removal of token economy and a return to baseline in the second A phase resulted in marked decreases in both points earned and behavioral ratings. After further replication on other patients, the authors concluded that institution of token economy effected positive changes in this type of depression. It should be underscored that it is only the obtained reversal in the second A phase that permits a firm conclusion with respect to the controlling effects of token economy on the two target behaviors. The reversal in the second A phase confirms that changes obtained in target behaviors were a

direct function of the institution and removal of the Token Economy treatment variable.

Although the A-B-A design is superior to the simple A-B design in that it permits unequivocal conclusions on the basis of the experimental analysis, two major problems are present. First, therapy for the patient who is being simultaneously treated and evaluated in this paradigm ends on the A or no-treatment phase of study. This was the case in the Hersen et al¹¹ study. One of the patients was discharged prematurely at his own request. For the other two, changes in medication necessitated termination of the experiment on the A phase, but clinical treatment was then completed. On an ethical and moral basis, it certainly behooves the experimenter-clinician to continue some form of treatment to its ultimate conclusion subsequent to completion of the research aspects of the case. A further design, known as the A-B-A-B design, meets this criticism as study ends on the B or treatment phase. This design has the added advantage of providing two occasions

for the treatment variable to demonstrate a positive effect, thus further strengthening conclusions as to its actions.

A-B-A-B Design

Miller¹² used an A-B-A-B experimental single-case design in an analysis of retention control training (RCT) in two "secondary enuretic" children. The number of enuretic episodes and mean frequency of daily urination were selected as target behaviors. During baseline (A) the patient was instructed to record target behaviors (Fig 3). He also received weekly treatment sessions consisting of discussion of troublesome situations that had occurred in the previous week. RCT (B) consisted of initially training the patient to refrain and postpone urination for ten minutes each time the urge was experienced. This was increased to 20 and 30 minutes in the following weeks. In addition, the patient was instructed to increase consumption of fluids throughout each day. During the second baseline (A) the patient was instructed to discontinue RCT, and in the following B phase it was reinstated. An examination of the data indicate that stable measurements were obtained in baseline. Application of RCT resulted in marked decreases in target behaviors. Removal of RCT in the second baseline led to both increased enuretic episodes and increased frequency of daily urination. Original baseline levels were achieved. Reintroduction of RCT led to decreased frequency of urination and eventual absence of enuretic episodes. This treatment was then continued until enuresis was virtually eliminated. As previously noted, this represents one of the major advantages over the less complete and less complex A-B-A design. In the Miller¹² experiment a double reversal in the data was obtained, indicating that the controlling effects of RCT could be replicated within the same patient, lending further credence to the effects of this treatment. An important experimental strategy to consider at this point is the length of the phases. Each phase should contain a relatively equal number of observations to insure that effects are due to the treatment variable. For instance, in the Miller¹² experiment enuresis could have worsened slightly at the very beginning of the B phase and then improved rapidly. Enuresis might have worsened somewhat again when the A phase was reintroduced. If the A phase was then stopped shortly thereafter, it could not be deter-

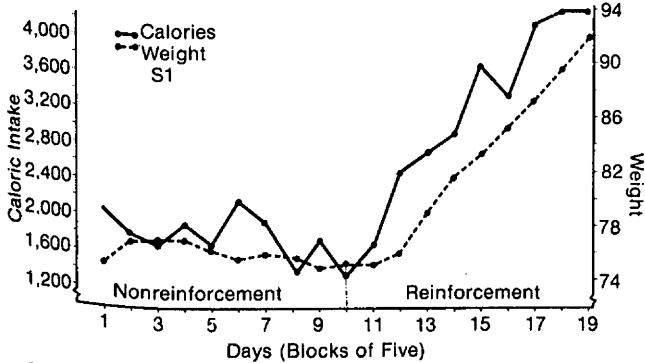


Fig 1.—Effects of nonreinforcement and reinforcement in a case of anorexia nervosa for subject 1. (From Leitenberg H et al¹⁰) reprinted by permission of the authors, editor, and publisher.)

Fig 2.—Number of points earned and mean behavioral ratings for subject 1. (From Hersen M et al¹¹; reprinted by permission of the authors, editor, and publisher.)

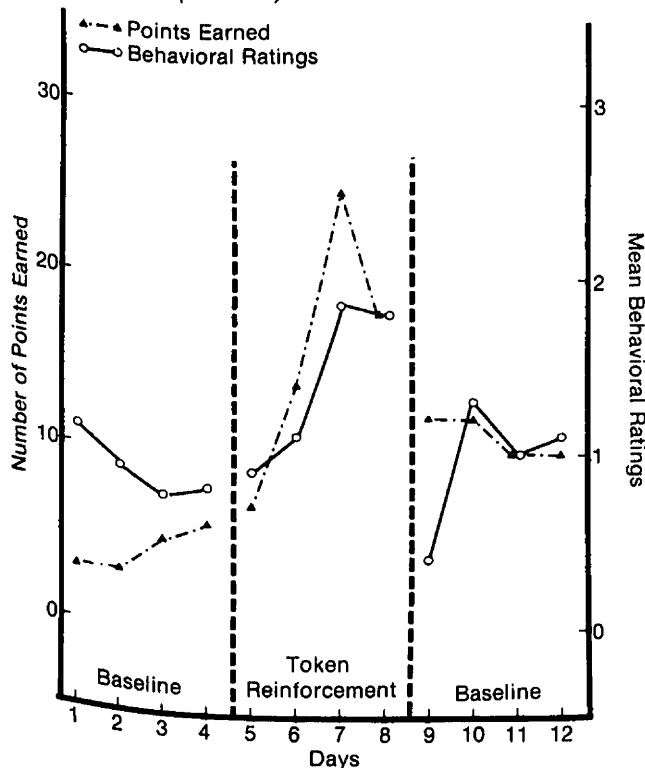
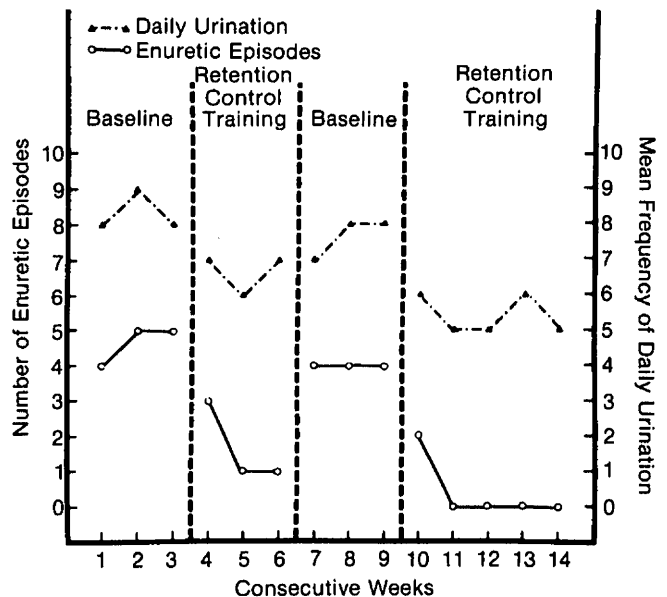


Fig 3.—Number of enuretic episodes per week and mean number of daily urinations per week for subject 1. (From Miller PM¹²; Behav Ther, reprinted by permission of the author, editor, and publisher.)



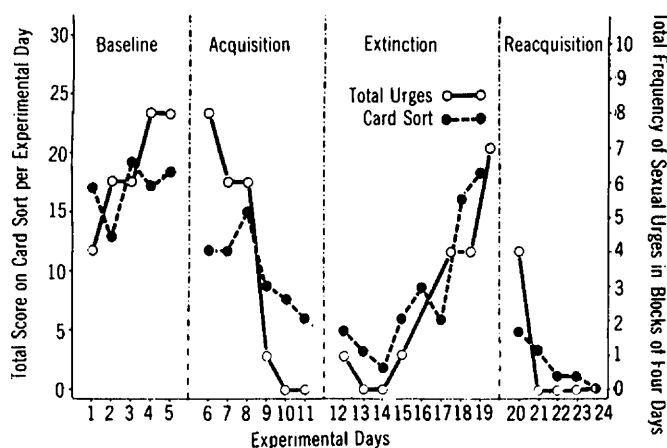


Fig 4.—Total score on card sort per experimental day and total frequency of pedophilic sexual urges in blocks of four days surrounding each experimental day. (Lower scores indicate less sexual arousal.) (From Barlow DH et al¹⁸; copyright 1969 by the American Psychological Association, and reproduced by permission.)

mined that the return to baseline produced the reversal. Such a reversal might have been due to some correlated event occurring at the change of treatment. Moreover, enuresis would have improved if the A phase had been extended to equal the previous B phase. In all single-case experiments, phases should be as nearly equal as possible.

B-A-B Design

Some researchers have used a variant of the A-B-A-B design, labeled the B-A-B design, in which the initial baseline phase is omitted.¹³⁻¹⁵ Others include an initial but abbreviated A phase in which only one data point is obtained.^{16,17} In both cases these designs are superior to the A-B-A design as study ends on the treatment or B variable. However, these designs do not allow for initial baseline assessment (A), and, consequently, it is not possible to determine changes brought about as a result of first introducing the B therapeutic variable. From both a research and clinical standpoint the use of the complete A-B-A-B design is recommended where possible for examination of the effects of singular therapeutic variables on behavior.

A-BC-B-BC Design

When the controlling effects of specific aspects of treatment techniques (eg, use of the noxious scene in covert sensitization) are to be examined, the experimental single-case design of choice is the A-BC-B-BC design.¹⁸ The A-BC-B-BC design is structurally similar to the A-B-A-B design, but procedurally different. As in the A-B-A-B design, the first phase of the A-BC-B-BC design involves a baseline assessment of target behaviors. In the BC phase a composite treatment variable is introduced, and changes in target behaviors, if any, are recorded and plotted graphically. In the B phase one aspect of the treatment variable is omitted in order to assess its controlling effects over target behaviors. If indeed that portion of the technique is critical for therapeutic success, improvement in target behaviors should cease or be reversed. By contrast, when that portion of the treatment variable is reintroduced in the BC phase, a second reversal should be obtained as indicated by renewed improvement in target behaviors.

A clear example of the A-BC-B-BC design is presented by Barlow et al¹⁸ in their assessment of the effects of the noxious scene in covert sensitization (a form of imaginal aversion therapy often used in treatment of sexual deviation and addictions). This was examined in one case of pedophilia and another of homosexuality.

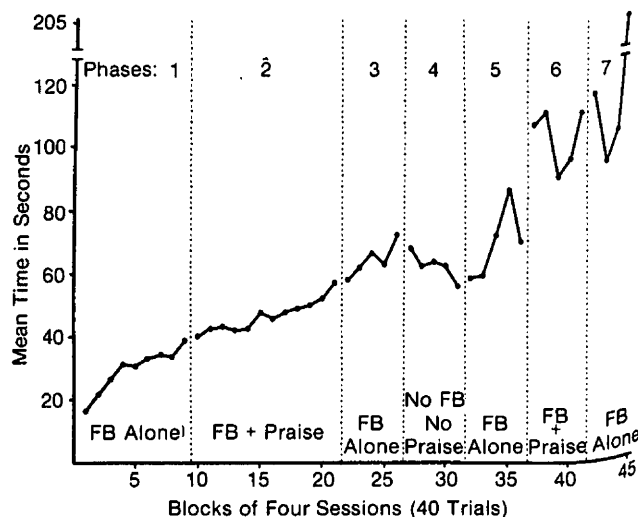


Fig 5.—Time in which a knife was kept exposed by a phobic patient as a function of feedback (FB), feedback plus praise, and no feedback or praise conditions. (Fig 2, p 136, from Leitenberg H et al¹⁴; copyright 1968 by Society for the Experimental Analysis of Behavior, Inc., and reproduced by permission.)

Two target behaviors were selected for study. One consisted of the patient's daily responses to a card sort containing a hierarchy of sexually arousing scenes. The number of daily urges toward "immature girls" was the second target behavior for the pedophilic patient. During baseline (A) operant rates of these target behaviors were recorded. Examination of Fig 4 shows an upward trend in both total urges and card sort scores during baseline (A). Covert sensitization procedures were then applied in acquisition (BC). Treatment involved daily imaginal presentation of deviant sexually arousing scenes paired with verbal descriptions of nausea and vomiting by the therapist. Examination of the figure indicates a dramatic linear decrease of urges and card sort scores during acquisition (BC). The specific or controlling effects of the noxious scene were then examined in extinction (B) for both total urges and card sort. Reintroduction of the noxious scene in reacquisition (BC) led to a renewed decrease of total urges and card sort scores, thus illustrating the *direct controlling effects* of pairing the noxious scene with the sexually arousing scene in covert sensitization.

It might be noted that a relatively equal number of data points are present in each experimental phase of the Barlow et al¹⁸ study. Secondly, it should be underscored that after introduction of treatment only one variable at a time was altered from one phase to the next (eg, elimination of the noxious scene in extinction while maintaining all other procedures as constants; reintroducing the noxious scene in reacquisition). If, on the other hand, two variables were to be manipulated simultaneously, it would not be possible to ascertain which of the two accounted for changes in the target behavior. Changing one variable at a time across conditions is an important guide rule in carrying out *all* types of experimental single-case research. Moreover, this rule is of particular importance when several therapeutic variables are simultaneously present.

Combined Designs

Combined designs have been used by a number of researchers during the course of their attempts to assess additive effects of particular psychotherapeutic variables on target behaviors.^{14,19-21} More specifically, if one therapeutic variable such as feedback is shown to effect changes in a target behavior, an experimental question might be raised as to the additional effects of a second therapeutic variable (eg, reinforcement) on that same behavior. The additive effects of the aforementioned two variables were examined by Leitenberg et al¹⁴ in their experimental treatment of a knife-phobic patient. The target behavior selected for study was the amount of time (in seconds) that the patient was able to re-

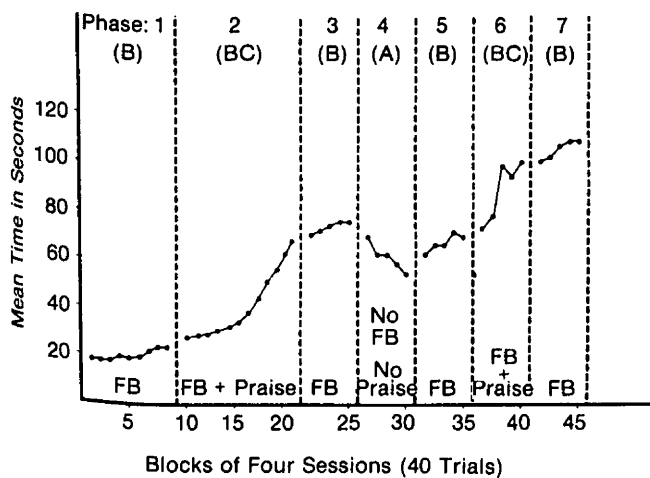


Fig 6.—Time in which a knife was kept exposed by a phobic patient as a function of feedback, feedback plus praise, and no feedback or praise conditions. (Hypothetical data based on Fig 2, p 136, from Leitenberg H et al.¹⁴)

main in the presence of the phobic object behind a closed door. The type of design used for this experimental analysis was a B-BC-B-A-B-BC-B design. B represented the feedback variable, BC a combination of feedback and praise, while A was baseline. Feedback consisted of informing the patient after each trial as to amount of time spent in all ten trials. Praise consisted of verbal reinforcement whenever the patient exceeded a progressively increasing time criterion. The results of this study are presented in Fig 5. During feedback (B) a marked upward linear trend was noted in the data. The addition of praise in phase 2 (BC) did not change the upward trend. Removal of praise in phase 3 (B) did not yield a change in the slope of the curve, suggesting that feedback alone was the critical variable controlling change. In phase 4 (A) a short reversal was obtained, confirming the controlling effects of the feedback alone variable. Reintroduction of feedback in phase 5 (B) led to renewed improvement. The addition of praise in phase 6 (BC) resulted in a continued upward trend. However, removal of praise in phase 7 (B) once again failed to bring about a change in the slope of the curve. In short, changes in the target behavior were purely a function of feedback alone. Praise did not produce an additive effect inasmuch as its controlling effects were not demonstrated in the experimental analysis. For purposes of comparison and illustration, data from the Leitenberg et al¹⁴ study are replotted to demonstrate the shape of the graph had the praise variable effected a controlling influence on the target behavior. It will be noted that in our hypothetical data in Fig 6 that a slight upward trend was obtained in phase 1 (B). The addition of praise in phase 2 (BC) resulted in a steep increase and a change in the slope of the curve. Removal of praise in phase 3 (B) led to a continued but less marked increase. Removal of feedback and praise in phase 4 (A) resulted in a reversal of the data while reinstatement of feedback in phase 5 (B) effected a slight upward trend. Addition of praise in phase 6 (BC) once again resulted in a steep upward trend and change in the slope of the curve. Removal of praise in phase 7 (B) resulted in a continued but slightly upward trend. These hypothetical data suggest that both therapeutic variables were effective. Feedback led to slight changes but marked changes in the slope of the curve were noted when praise was added, thus illustrating both additive and controlling effects of that variable.

In some cases, as many as three and four variables have been examined sequentially and in combination.^{19,20} Many variants of the combined design are possible; three examples are A-B-BC-B-BC, BC-B-BC-B, and A-B-BC-BCD-BC-BCD.

Multiple Baseline Designs

In all of the experimental analysis designs presented to this point, the reversal technique (removal of the treatment variable

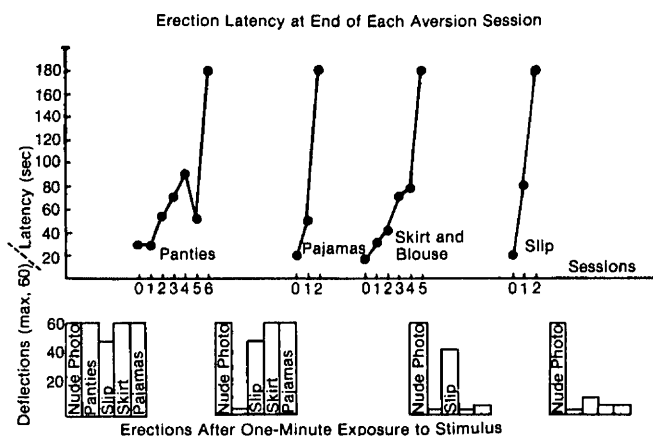


Fig 7.—Specificity of autonomic changes (patient B). (From Marks IM and Gelder MG²²; reprinted by permission of the authors and editor.)

or a relevant portion of it) has been used to demonstrate the controlling effects of the therapeutic variable under study. There are times, however, when the reversal technique may not be feasible, particularly when treatment considerations argue against its application. An alternative method to demonstrate the controlling effects of therapeutic variables in single subjects is known as the "multiple baseline" technique.²¹ In this method of study *specific but independent* target behaviors are identified and precisely defined. A baseline measurement of each target behavior is established, following which a particular therapeutic technique (eg, electrical aversion) is applied to the first of the target behaviors. If the technique is successful and the selected target behaviors are truly independent of one another, changes in the first target behavior should appear while little or no change is noted in the others. Subsequently, the technique is applied to a second target behavior, and changes are again noted. However, such changes should not be found in the remaining untreated behaviors. Baer et al²² argue that: "The experimenter is attempting to show that he has a reliable experimental variable, in that each behavior changes maximally only when the experimental variable is applied to it." In application, the "multiple baseline" design is completed when the therapeutic variable has been administered to each of the designated target responses. There are no specific rules as to how many target behaviors are needed to establish control and specificity of treatment, but the controlling effects of that technique over at least three target behaviors would appear to be a minimum requirement.

The "multiple baseline" design was used by Marks and Gelder²² in their assessment of electrical aversion therapy in treating sexual deviation. One of their study patients was a young male transvestite. Baseline assessment indicated that sexual arousal (measured via a penile transducer) was maximal when he either observed or touched one of several stimuli (panties, slip, skirt, woman's pajamas) for a period of one minute. All of these stimuli had previously been used in his cross-dressing episodes. Similarly, the patient responded maximally to a photograph of a nude female. Following baseline assessment, a course of electrical aversion consisting of about 20 trials was administered to the patient in relation to each of the target stimuli (panties, slip, skirt, woman's pajamas) in sequence. Erection latency to each target stimulus following aversion sessions is presented at the top of Fig 7. Strength of penile erection after a one-minute exposure to target stimuli is presented at the bottom of Fig 7. It will be noted that erectile strength (first block of five stimuli) prior to aversion

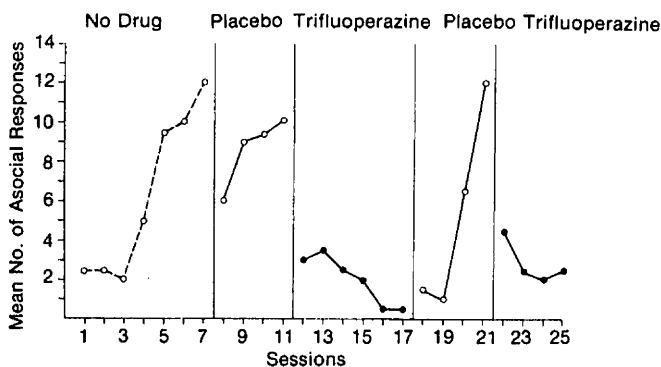


Fig 8.—Average number of refusals to engage in a brief conversation during 18 random time samples per day by patient. Trifluoperazine (Stelazine) dose was 60 mg daily. Each session represents the average of a two-day block of observation. (From Liberman RP, et al: Research design for analysing drug-environment-behavior interactions. *J Nerv Ment Dis*,²⁸ reprinted by permission of the authors and the Williams & Wilkins Co, Baltimore.)

was maximal, with the exception of the "slip" stimulus. The second block of five stimuli shows erectile strength following electrical aversion in connection with the panties stimulus. Erectile strength in response to panties was extremely low, but it was at maximum strength for the other four stimuli. The third block of five stimuli depict erectile strength following sequential aversion with respect to panties, skirt, and woman's pajamas. Examination of that block reveals maximum erectile strength to the nude photo and marked erectile strength to the untreated "slip" stimulus. Aversion was then applied in relation to the "slip" stimulus, with a resulting decrease in erectile strength (fourth block). Erectile strength towards the nude photo remained at its maximum, indicating the specificity of aversion treatment to these stimuli. In this study erectile responses to the nude photo may be conceptualized as a control for the other four stimuli. If erections to untreated articles of clothing and the nude photo had decreased, then one would either conclude that electrical aversion had some general effect on sexual arousal or that a correlated therapeutic variable (eg, expectancy) was the crucial therapeutic agent. In summary, Marks and Gelder²² demonstrated, both sequentially and differentially, the controlling effects of electrical aversion over erectile responses to five target stimuli selected for study.

Side Effects

From the foregoing description of single-case research, it becomes apparent that a wide variety of problem areas have been studied. However, there are still other features and applications of these designs that have not been discussed. For example, some investigators²³⁻²⁷ have not only examined the controlling effects of particular variables on designated target behaviors, but have also measured and assessed the "side effects" of these same variables on other ongoing nonmanipulated behaviors. Sajwaj et al²⁸ point out that such covariations in nonmanipulated behaviors may possibly result in both socially desirable and undesirable side effects. This is of particular importance in the case of undesirable side effects, as additional application of techniques will then be required to exert appropriate controls over behaviors.

Effects of Drugs

The experimental single-case design is also well suited for examination of the effects of pharmacological agents on behavior.² Using experimental analysis designs under double-blind conditions, one might sequentially administer a placebo, drug X, a placebo, and once again drug X (A-B-A-B) while observing concomi-

tant behavioral changes. One might also examine two drugs in sequence (placebo, drug X, placebo, drug Y, placebo, drug X, placebo, drug Y), and in other instances the additive effects of drugs (placebo, drug X, drug X and drug Y, drug X, drug X and drug Y). Since continued measurements are in effect, length of phases can be varied from experiment to experiment to determine precisely the latency of drug effects after beginning the dosage and the residual effects after discontinuing the dosage.

An example of the effects of a drug on a selected target behavior in a within-subject reversal design is presented by Liberman et al.²⁸ The effects of trifluoperazine (Stelazine) were assessed in an A-A1-B-A1-B design in a 21-year-old withdrawn male schizophrenic. The target behavior chosen for study was the patient's willingness to engage in five-minute chats initiated by a member of the nursing staff (blind to conditions) 18 times daily at randomly selected times. On day 1 the patient was withdrawn from his trifluoperazine medication and an examination of Fig 8 indicates that during the A phase (no drug) the number of asocial responses (unwillingness to chat) increased sharply. Institution of the placebo in the A1 phase (placebo) resulted in an initial decrease followed by a marked linear increase in asocial responses. In phase B (trifluoperazine) a dosage of 60 mg/day was introduced. A marked decrease in asocial responses resulted. When placebo was reintroduced in the A1 phase (placebo), a reversal was noted as indicated by the marked increase in asocial responses. Reintroduction of trifluoperazine (60 mg/day) in the second B phase led to a second reversal (decrease in asocial responses), thus suggesting the controlling effects of the drug. The conventional double-blind design used in group drug studies was not quite applicable here. Although the patient was unaware of placebo conditions, the experimenter (physician) was obviously aware of the drug being administered. However, the assessor of the target behavior (eg, the nurse) was blind to the condition in force. In that sense the spirit of the double-blind design is approximated and maintained.

Comment

Obviously the single-case design cannot answer all clinical research questions. A first limitation occurs if one wishes to compare two global treatments, both of which are effective on a given behavior disorder. To find out which treatment is more effective would be difficult to answer using single-case methodology. In this case a group design, where each treatment is administered to a separate set of patients and where results are analyzed statistically, would be more appropriate.

One must question, however, the usefulness of determining which treatment is statistically better when differences, unless extraordinarily large by statistical standards, are clinically unimportant. This issue has been raised by Bergin and Strupp¹ who, in a thorough review of psychotherapy research, concluded that any effort designed to evaluate global treatments such as "psychotherapy" is likely to produce "weak" (clinically insignificant) results. A more productive approach at this stage of our development of behavioral change techniques would be to determine through single-case designs those active ingredients that are effective in both treatments. These ingredients could then be combined into a more effective composite therapy.

A second limitation of single-case designs arises if one wishes to test variables that are irreversible or partially irreversible (eg, the effects of surgical lesions or certain types of therapeutic instructions). Here the effect of the therapeutic variable continues in subsequent phases, pro-

ducing what is termed "carry over effects" or "sequential confounding of results." These therapeutic variables may, at times, be tested in multiple baseline designs where several target behaviors are independently measured, but often a group comparison is the only alternative. (Note, however, in some instances therapeutic instructions can be tested in single-case designs.²⁰)

A similar difficulty ensues if one is testing a therapeutic variable with long-lasting effects such as some pharmacologic agents. Here one must balance the disadvantages of alternating the drug and a placebo in phases each lasting several months with the previously mentioned disadvantages of the group comparison.

Finally, one of the most important aspects of research is the generality of the conclusions. One of the assumed strengths of group designs is that results are applicable to all patients carrying the same diagnosis. The fallacies of this assumption were discussed in the introduction where it was noted that a few people may improve a great deal and many deteriorate somewhat, producing an overall statistical improvement but losing important individual differences in the group average. Although single-case designs provide clinicians with information on effects of

treatment for patients with the same characteristics, their results also may not be applicable to all patients with a similar behavior disorder. The answer to the problem is to apply *systematically* the therapeutic variable to cases with different background behaviors or "personality" variables, an approach which Sidman³ refers to as "systematic replication." Failures with a systematic replication series can then be ascribed to specific patient characteristics. Experimental analyses should then be performed to determine necessary alterations in therapeutic procedures. This strategy highlights individual differences rather than averaging them out.

This present review is by no means an exhaustive account of single-case experimental designs or strategy. However, suitability of this approach to clinical research should lead to many variations of these designs as we strive to answer complex questions concerning the treatment of human behavior disorders.

Preparation of this manuscript study was supported in part by grant MH-20258 from the National Institute of Mental Health and by Veterans Administration grant 5-71 (Jackson Veterans Administration Center, Jackson, Miss).

References

1. Bergin AE, Strupp HH: *Changing Frontiers in the Science of Psychotherapy*. Chicago, Aldine-Atherton, 1972.
2. Shapiro MB: The single case in fundamental clinical psychological research. *Br J Med Psychol* 34:255-262, 1961.
3. Sidman M: *Tactics of Scientific Research*. New York, Basic Books Inc Publishers, 1960.
4. Dukes WF: N-1. *Psych Bull* 64:74-79, 1965.
5. Chassan JB: *Research Design in Clinical Psychology and Psychiatry*. New York, Appleton-Century-Crofts, 1967.
6. Baer DM, Wolf MM, Risley TR: Some current dimensions of applied behavior analysis. *J Appl Behav Anal* 1:91-97, 1968.
7. Davidson PO, Costello CG (eds): *N-1: Experimental Studies of Single Cases*. New York, Van Nostrand-Reinhold Co, 1969.
8. Bergin AE: Some implications of psychotherapy research for therapeutic practice. *Int J Psychiatry* 3:136-150, 1967.
9. Gentile JR, Roden AH, Klein RD: An analysis-of-variance model for the intrasubject replication design. *J Appl Behav Anal* 5:193-198, 1972.
10. Leitenberg H, Agras WS, Thomson L: A sequential analysis of the effect of selective positive reinforcement in modifying anorexia nervosa. *Behav Res Ther* 6:211-218, 1968.
11. Hersen M, et al: Effects of token economy on neurotic depression: An experimental analysis. *Behav Ther*, to be published.
12. Miller PM: An experimental analysis of retention control training in the treatment of nocturnal enuresis in two institutionalized adolescents. *Behav Ther*, to be published.
13. Ayllon T, Azrin NH: The measurement and reinforcement of behavior of psychotics. *J Exp Anal Behav* 8:357-383, 1965.
14. Leitenberg H, et al: Feedback in behavior modification: An experimental analysis in two phobic cases. *J Appl Behav Anal* 1:131-137, 1968.
15. Rickard HC, Saunders TR: Control of "clean-up" behavior in a summer camp. *Behav Ther* 2:340-344, 1971.
16. Agras WS, Leitenberg H, Barlow DH: Social reinforcement in the modification of agoraphobia. *Arch Gen Psychiatry* 19:423-427, 1968.
17. Whitman TL, Zakaras M, Chardos S: Effects of reinforcement and guidance procedures on instruction-following behavior of severely retarded children. *J Appl Behav Anal* 4:283-290, 1971.
18. Barlow DH, Leitenberg H, Agras WS: Experimental control of sexual deviation through manipulation of the noxious scene in covert sensitization. *J Abnorm Psychol* 5:596-601, 1969.
19. Agras WS, et al: Instructions and reinforcement in the modification of neurotic behavior. *Am J Psychiatry* 125:1435-1439, 1969.
20. Elkin TE, et al: Modification of caloric intake in anorexia nervosa: An experimental analysis. *Psychol Rep* 32:75-78, 1973.
21. Hersen M, et al: Instructions and reinforcement in the modification of a conversion reaction. *Psychol Rep* 31:719-722, 1972.
22. Marks IM, Gelder MG: Transvestism and fetishism: Clinical and psychological changes during faradic aversion. *Br J Psychiatry* 113:711-729, 1967.
23. Lovaas OI, Simmons JQ: Manipulation of self-destruction in three retarded children. *J Appl Behav Anal* 2:143-158, 1969.
24. Risley TR: The effects and side-effects of punishing the autistic behavior of a deviant child. *J Appl Behav Anal* 1:21-34, 1968.
25. Sajwaj T, Twardosz S, Burke M: Side effects of extinction procedures in a remedial preschool. *J Appl Behav Anal* 5:163-175, 1972.
26. Twardosz S, Sajwaj T: Multiple effects of a procedure to increase sitting in a hyperactive retarded boy. *J Appl Behav Anal* 5:73-78, 1972.
27. Wahler RG, et al: The modification of childhood stuttering: Some response-response relationships. *J Exp Child Psychol* 9:411-428, 1970.
28. Liberman RP, et al: Research design for analyzing drug-environment-behavior interactions. *J Nerv Ment Dis*, to be published.
29. Barlow DH, et al: The contribution of therapeutic instruction to covert sensitization. *Behav Res Ther* 10:411-415, 1972.